

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

JCS11 U.S. PTO
09/309894
05/11/99

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1998年10月29日

出 願 番 号

Application Number:

平成10年特許願第308943号

出 願 人

Applicant (s):

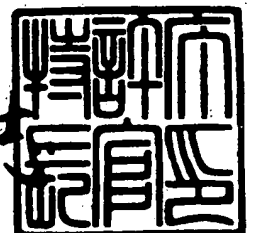
富士通株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

1999年 3月12日

特 許 庁 長 官
Commissioner,
Patent Office

伴 佐 山 建 志



出証番号 出証特平11-3014866

【書類名】 特許願

【整理番号】 9804384

【提出日】 平成10年10月29日

【あて先】 特許庁長官殿

【国際特許分類】 G06K 9/62

【発明の名称】 単語認識装置および方法

【請求項の数】 13

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 武部 浩明

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 堀田 悦伸

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 直井 聡

【特許出願人】

 【識別番号】 000005223

 【氏名又は名称】 富士通株式会社

【代理人】

 【識別番号】 100074099

 【郵便番号】 102

 【住所又は居所】 東京都千代田区二番町8番地20 二番町ビル3F

 【弁理士】

 【氏名又は名称】 大菅 義之

 【電話番号】 03-3238-0031

【選任した代理人】

【識別番号】 100067987

【郵便番号】 222

【住所又は居所】 神奈川県横浜市港北区太尾町 1418-305（大倉
山二番館）

【弁理士】

【氏名又は名称】 久木元 彰

【電話番号】 045-545-9280

【手数料の表示】

【予納台帳番号】 012542

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9705047

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 単語認識装置および方法

【特許請求の範囲】

【請求項 1】 1 つ以上の単語のリストを格納するリスト手段と、
複数の文字の特徴量を格納する辞書手段と、
前記辞書手段に格納された前記文字の特徴量を用いて、前記リスト手段に格納された単語の特徴量を生成する生成手段と、
生成された単語の特徴量を認識対象の特徴量と照合し、認識結果を出力する照合手段と
を備えることを特徴とする単語認識装置。

【請求項 2】 前記照合手段は、前記単語の特徴量を格納するメモリ手段を含み、該単語の特徴量の照合が終了したとき、該メモリ手段を解放して、次の単語の特徴量を格納することを特徴とする請求項 1 記載の単語認識装置。

【請求項 3】 前記認識対象として画像を入力する入力手段と、入力された画像の各小領域における輪郭線方向の方向指数ヒストグラムに対して文字の連結方向と垂直な方向に 1 次元のぼけ変換を施し、変換結果から得られた方向指数ヒストグラム系列を前記認識対象の特徴量として抽出する抽出手段とをさらに備えることを特徴とする請求項 1 記載の単語認識装置。

【請求項 4】 前記抽出手段は、前記文字の連結方向と垂直な方向における前記入力された画像の長さをあらかじめ決められた整数で割り、得られた商を前記小領域の大きさとして用いて、該画像を各小領域に分割することを特徴とする請求項 3 記載の単語認識装置。

【請求項 5】 前記生成手段は、複数の文字の特徴量を合成して前記単語の特徴量を生成することを特徴とする請求項 1 記載の単語認識装置。

【請求項 6】 前記生成手段は、前記単語を構成する各文字の特徴量に対応する方向指数ヒストグラム系列を並べて新たな方向指数ヒストグラム系列を生成し、生成された方向指数ヒストグラム系列を前記単語の特徴量とすることを特徴とする請求項 5 記載の単語認識装置。

【請求項 7】 前記照合手段は、前記単語の特徴量と前記認識対象の特徴量

の非線形マッチングを行い、該単語の特徴量と該認識対象の特徴量の類似度を計算することを特徴とする請求項 1 記載の単語認識装置。

【請求項 8】 前記リスト手段は、前記認識対象に対応する単語を含んでいる可能性の高いリストを格納することを特徴とする請求項 1 記載の単語認識装置。

【請求項 9】 文字の特徴量を用いて単語の特徴量を動的に生成する生成手段と、

生成された単語の特徴量を認識対象の特徴量と照合し、認識結果を出力する照合手段と

を備えることを特徴とする単語認識装置。

【請求項 10】 パターンの特徴量を用いてパターン列の特徴量を動的に生成する生成手段と、

生成されたパターン列の特徴量を認識対象の特徴量と照合し、認識結果を出力する照合手段と

を備えることを特徴とする認識装置。

【請求項 11】 コンピュータのためのプログラムを記録した記録媒体であって、

文字の特徴量を用いて単語の特徴量を動的に生成するステップと、

生成された単語の特徴量を認識対象の特徴量と照合するステップと

を含む処理を前記コンピュータに実行させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項 12】 コンピュータのためのプログラムを記録した記録媒体であって、

パターンの特徴量を用いてパターン列の特徴量を動的に生成するステップと、

生成されたパターン列の特徴量を認識対象の特徴量と照合するステップと

を含む処理を前記コンピュータに実行させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項 13】 1 つ以上のパターン列のリストを生成し、

複数のパターンの特徴量を格納する辞書を生成し、

前記辞書に格納された前記パターンの特徴量を用いて、前記リストに格納されたパターン列の特徴量を動的に生成し、

生成されたパターン列の特徴量を認識対象の特徴量と照合することを特徴とする認識方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、パターン列を一括して認識するパターン認識に係り、特に、単語を一括して認識する単語認識装置とその方法に関する。

【0002】

【従来の技術および発明が解決しようとする課題】

従来の単語認識方法は、文字切り出しの観点から次の3つの方法に分類できる。

【0003】

第1の方法は、画像特徴を利用して単語を1文字単位に切り出し、切り出された文字を個別に文字認識する方法である。主な画像特徴としては、文字間の空白やピッチ、画像を文字列方向と垂直な方向に射影して得られるヒストグラム、画素の連結成分の外接矩形、画像の上部輪郭および下部輪郭の凹凸等がある。

【0004】

第2の方法は、切り出し仮説を複数個立て、それぞれの仮説を文字認識結果によって検定する方法である。切り出し仮説は、観測窓を画像内で動かして得る場合と、上述の画像特徴を利用して得る場合とがある。また、検定には、全体の整合性をとるために、動的計画法 (Dynamic Programming, DP) を用いることが多い。

【0005】

しかしながら、制約のない状態で書かれた手書き文字列の場合、文字間ピッチは一定ではなく、切断すべき部分の画像特徴が極めて多様であるため、切り出しがうまくいかないという問題がある。観測窓でサーチする場合でも、文字の大きさやピッチが一定でないため、固定の窓では処理しきれないし、窓の大きさを可

変にすれば、処理時間が膨大になる。

【0006】

さらに、切断すべき部分の画像特徴は、漢字、ひらがな、英字、数字のそれぞれの字種に固有であるから、印刷文字の単語においても、それらの字種が混合した場合、接触文字分離の際に同様の問題が生じる。

【0007】

第3の方法は、単語を文字単位に切り出すことなく、単語そのものを一括して認識する方法である。この方法は、文字の切り出しという困難な問題を回避できるものの、個別文字を認識する場合に比べて、認識辞書に登録すべき候補の数が飛躍的に増加するという問題がある。実際には、メモリの量的制約に基づいて認識辞書のサイズを実用的な範囲に抑えるために、限られた個数内の単語しか認識辞書に登録できず、用途が制限される。

【0008】

本発明の課題は、認識可能な単語の範囲をできるだけ制限することなく、単語を一括して認識する単語認識装置とその方法を提供することである。

【0009】

【課題を解決するための手段】

図1は、本発明の単語認識装置の原理図である。図1の単語認識装置は、リスト手段1、辞書手段2、生成手段3、および照合手段4を備える。

【0010】

リスト手段1は、1つ以上の単語のリストを格納し、辞書手段2は、複数の文字の特徴量を格納する。生成手段3は、辞書手段2に格納された文字の特徴量を用いて、リスト手段1に格納された単語の特徴量を生成する。そして、照合手段4は、生成された単語の特徴量を認識対象の特徴量と照合し、認識結果を出力する。

【0011】

リスト手段1が格納するリストには、認識結果の候補となる単語が登録されており、辞書手段2には、それらの単語を構成する個別文字の特徴量が登録されている。生成手段3は、リスト手段1のリストを参照し、1つの単語を構成する各

文字の特徴量を辞書手段 2 から取り出して、その単語の特徴量を合成する。照合手段 4 は、生成手段 3 が合成した単語の特徴量を、入力画像に含まれる認識対象の特徴量と照合し、2 つの特徴量の類似度等を求めて、それを認識結果として出力する。

【0012】

このように、あらかじめ多数の単語の特徴量を登録した単語辞書を用意するのではなく、認識処理の過程で候補となる単語のみについて特徴量を動的に生成することで、使用するメモリ量を削減することができる。また、リストには単語の特徴量を登録する必要がないので、多数の単語を登録することができ、それらの単語の特徴量を随時生成することができる。このため、従来の単語辞書のように、単語の範囲が制限されることはない。

【0013】

例えば、図 1 のリスト手段 1、辞書手段 2、生成手段 3、および照合手段 4 は、それぞれ、後述する図 2 の単語リスト 14、個別文字辞書 15、特徴生成部 13、および特徴照合部 12 に対応する。

【0014】

【発明の実施の形態】

以下、図面を参照しながら、本発明の実施の形態を詳細に説明する。以下の説明では、簡潔な記述を行うため、認識対象の画像に含まれる文字列が横書きである場合を想定している。したがって、文字の連結方向は横方向であり、単語も横書きである。しかしながら、本発明は、文字列が縦書きで文字の連結方向が縦方向である場合にも、同様に適用することができる。

【0015】

本実施形態の単語認識装置は、個別文字の辞書から動的に単語辞書を生成し、単語を一括認識する。このような単語認識装置を実現する際のキーとなるポイントは、単語画像から得られる特徴量と単語を構成する各文字の特徴量の合成結果とが一致するように、特徴量および特徴量の合成方法を決定することである。

【0016】

個別文字の辞書には、各文字の特徴量が登録されており、複数の文字の特徴量

を合成することにより、対応する単語の特徴量が生成される。今、文字Aと文字Bとからなる単語X(=“AB”)を考え、文字A、文字B、および単語Xの画像をそれぞれ画像a、画像b、画像xとし、それらの画像から得られる特徴量をそれぞれ α 、 β 、 χ とする。

【0017】

このとき、個々の文字の特徴量から単語の特徴量を生成するためには、2つの特徴量間にある合成演算 f が定義され、 $\chi = f(\alpha, \beta)$ が成り立たなくてはならない。このような条件が成り立つことを、演算 f は画像の通常の合成に関して可換であるということにする。

【0018】

ところで、従来より用いられている文字および単語の特徴量には、文字位置のずれや字体の変形を救済するために、ぼかし処理が加えられている。ぼかし処理は、画像をある決められた個数でメッシュ分割し、得られた各小領域における方向指数(direction code)ヒストグラムを周囲の小領域のそれに重み付き加算する処理である。このぼかし処理により、1つの小領域の情報に周囲の小領域の情報が取り入れられるため、文字のずれや変形を吸収することができる。

【0019】

しかしながら、このようなぼかし処理を単語に適用した場合、2つの文字の境界で一方の文字の方向指数ヒストグラムが他方のそれに重み付き加算されるため、可換な合成演算 f は簡単には見つからない。

【0020】

そこで、本実施形態では、各小領域における方向指数ヒストグラムに対して、文字の連結方向に垂直な縦方向には従来と同様のガウス分布等によるぼかし処理を施し、文字の連結方向である横方向にはぼかし処理を施さないで、特徴量を生成する。

【0021】

横書き単語にこのような特徴量を用いた場合、2つの文字の境界で一方の文字の方向指数ヒストグラムが他方のそれに重み付き加算されることがなく、可換な合成演算 f は、単純に2つの文字の特徴量を並べる操作で与えられる。ただし、

このままでは横方向の文字のずれや変形が考慮されないため、画像と認識候補の間の距離計算時にDPマッチングを用いて認識精度の向上を図る。

【0022】

また、従来の文字の特徴量は画像を一定数のメッシュで分割することにより得られるが、同様のメッシュ分割を複数の文字が連なる単語にそのまま適用した場合、単語に含まれる文字数が増えれば、それだけ相対的にメッシュは粗くなる。このため、メッシュの解像度が相対的に低くなり、認識精度に悪影響を与える。

【0023】

そこで、本実施形態では、単語の長さに応じてメッシュの個数を可変にする。横書きの単語の場合、文字数が増えても、画像の縦の長さは大きく変わらないので、画像の縦の長さを一定数で割り、得られた商を基本メッシュの大きさと決定する。そして、その大きさに基づいて縦横にメッシュ分割を行う。この場合、画像の横の長さに応じて横方向のメッシュ数は変化するが、距離計算でDPマッチングを用いるので、メッシュ数の変化による不確定性は吸収される。

【0024】

縦書きの単語の特徴量については、上述した横書きの単語の特徴量の説明において、縦と横を入れ替えればよい。

図2は、本実施形態の単語認識装置の構成図である。図2の単語認識装置は、特徴抽出部11、特徴照合部12、特徴生成部13、単語リスト14、および個別文字辞書15を備える。

【0025】

特徴抽出部11は、与えられた画像から特徴量を抽出し、特徴生成部13は、個別文字辞書15に登録された文字の特徴量を用いて、単語リスト14に登録された認識候補の単語の特徴量を合成する。特徴照合部12は、特徴生成部13が生成した単語の特徴量を単語辞書として用いて、特徴抽出部11が抽出した特徴量を、各単語の特徴量と比較する。そして、最も類似した特徴量を持つ単語を認識結果の第1候補として出力する。

【0026】

このとき、単語リスト14は、認識対象の画像が表す単語を含んでいることが

望ましいが、単語の数が多過ぎると処理が煩雑になる可能性がある。そこで、あらかじめいくつかの単語リスト14を用意しておき、特徴生成部13は、直前の画像の認識結果に応じて、認識対象の単語を含んでいる可能性の高い単語リスト14を推定し、その単語リスト14を用いる。

【0027】

例えば、多数の手紙に書かれた住所の画像を処理する場合、直前の認識結果が“都”、“道”、“府”、または“県”で終わっていれば、次に市町村名が現れる可能性が高いと判断し、市町村名を含む単語リスト14を選択して、単語の特徴量を合成する。

【0028】

次に、図3から図9までを参照しながら、特徴抽出部11の処理について説明する。図3は、特徴抽出部11の処理のフローチャートである。この処理は、「加重方向指数ヒストグラム法による手書き漢字・ひらがな認識 (Handwritten “KANJI” and “HIRAGANA” Character Recognition Using Weighted Direction Index Histogram Method)」(鶴岡 信治ほか、電子情報通信学会論文誌 D, Vol. J70-D, No. 7, pp. 1390-1397, 1987年7月.)に記載された処理を参考にして、新たな変更を加えたものである。

【0029】

まず、特徴抽出部11は、認識対象の画像を入力し(ステップS1)、方向指数ベクトル場を作成する(ステップS2)。次に、画像の基本メッシュ分割を行って方向指数ヒストグラムベクトル場を作成し(ステップS3)、1次元ガウス分布関数による縦方向のぼけ変換を行う(ステップS4)。そして、方向指数ヒストグラムベクトル場のベクトルを圧縮して(ステップS5)、特徴量を抽出し、処理を終了する。

【0030】

ステップS2において、特徴抽出部11は、まず、入力画像について8連結の輪郭追跡を行い、その結果得られる輪郭点系列を $\{C_i\}$ とする。ここで、 C_i は、画像に含まれるパターンの輪郭上の画素に対応する。次に、輪郭点 C_i における8方向の方向指数 d_i を、その1つ後の輪郭点 C_{i+1} の位置に基づいて決

定する。

【0031】

図4は、輪郭点 C_i を中心とした場合の輪郭点 C_{i+1} の位置と方向指数の関係を示している。例えば、 C_{i+1} が C_i の右にあれば $d_i = 1$ となり、 C_{i+1} が C_i の右上にあれば $d_i = 3$ となり、 C_{i+1} が C_i の上にあれば $d_i = 5$ となる。

【0032】

次に、 C_i における方向指数 d_i と、 C_i の1つ前の輪郭点 C_{i-1} における方向指数 d_{i-1} との平均を取ることで、図5に示すような、 C_i における16方向の方向指数 D_i を得る。

【0033】

例えば、 C_{i-1} 、 C_i 、および C_{i+1} が図6に示すような位置関係にある場合、 C_i は C_{i-1} の下にあるので、 $d_{i-1} = 13$ となり、 C_{i+1} は C_i の左下にあるので、 $d_i = 11$ となる。したがって、16方向の方向指数 D_i は、 $((d_{i-1}) + d_i) / 2 = 12$ となる。この方向指数は、方向指数13の方向と方向指数11の方向の中間の方向を表している。一般に、方向指数 D_i が奇数の場合は、図4に示した8つの方向のいずれかを表し、偶数の場合は、それらのうち隣接する2つ方向の中間の方向を表す。

【0034】

次に、画像の各点（画素）に対して、16値ベクトルを割り当てる。ここでは、輪郭点以外の点には、16個の成分がすべて0のベクトルを割り当て、輪郭点 C_i には、 D_i 番目の成分が1でそれ以外の成分がすべて0のベクトルを割り当てる。これらの16値ベクトルからなるベクトル場は、方向指数ベクトル場と呼ばれる。

【0035】

例えば、図6の輪郭点 C_i における方向指数ベクトルは、

$(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0)$

となる。

【0036】

また、ステップ S 3 において、特徴抽出部 11 は、まず、画像の縦方向の長さ y をあらかじめ決められた整数 M で割り、その商 L を基本メッシュの大きさとする。そして、画像の横方向の長さ x を L で割った商を n として、画像全体を $M \times n$ のメッシュで分割する。このようなメッシュ分割によれば、横方向の長さに応じてメッシュの数が増減し、常に一定の大きさのメッシュを得ることができる。

【0037】

次に、得られた各メッシュにおいて、方向指数 D_i のヒストグラムを作成する。ただし、ヒストグラムの重み係数はすべて 1 とする。このヒストグラムは、メッシュ内に含まれる点の方向指数ベクトルを加算することにより生成され、16 値ベクトルで表される。そこで、これを方向指数ヒストグラムベクトルと呼ぶことにし、すべてのメッシュの方向指数ヒストグラムベクトルからなるベクトル場を、方向指数ヒストグラムベクトル場と呼ぶことにする。例えば、

(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)

(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0)

(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0)

(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0)

のような 4 つの方向指数ベクトルからは、次のような方向指数ヒストグラムベクトルが得られる。

(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0)

また、ステップ S 4 において、特徴抽出部 11 は、縦方向のみについて、1 次元ガウス分布関数を重み係数として用いたぼけ変換を行う。ここでは、例えば、図 7 に示すような 5 つの重み係数からなる 1 次元ガウス分布型フィルタを生成し、それを縦方向に並んだ 5 つのメッシュの方向指数ヒストグラムベクトルに適用する。

【0038】

これにより、5 つの方向指数ヒストグラムベクトルの各成分がガウス分布に従って重み付き加算され、新たな方向指数ヒストグラムベクトルが生成される。そして、中心位置のメッシュの方向指数ヒストグラムベクトルは、生成された方向指数ヒストグラムベクトルにより更新される。

【0039】

このように、1次元ガウス分布型フィルタを用いて縦方向にのみぼけ変換を行うことで、縦方向の文字のずれや変形を吸収することができる。一方、文字の連結方向である横方向については、方向指数ヒストグラムベクトルが重み付き加算されず、2つの文字の境界で特徴が混ざり合うことがない。したがって、上述の可換な合成演算 f を、後述するように簡単に定義することができる。

【0040】

ここで、 $m < M$ となるような整数 m をあらかじめ決めておき、縦方向の M 個のメッシュの中から m 個のメッシュをフィルタの中心位置として選べば、 $M \times n$ の方向指数ヒストグラムベクトル場を、 $m \times n$ の方向指数ヒストグラムベクトル場に空間情報圧縮することができる。

【0041】

例えば、 $M = 13$ および $n = 13$ の場合、画像は、図8に示すような 13×13 のメッシュに分割される。ここで、 $m = 7$ とし、○印の付いたメッシュをフィルタの中心位置として選べば、 13×13 の方向指数ヒストグラムベクトル場は 7×13 の方向指数ヒストグラムベクトル場に空間情報圧縮される。

【0042】

また、ステップ S5 において、特徴抽出部 11 は、方向指数ヒストグラムベクトル場のベクトルを、方向圧縮する。まず、方向指数ヒストグラムベクトルの 16 個の成分のうち、方向指数 1、3、5、7、9、11、13、および 15 に対応する各成分の値に、その前後の成分の値を 0.5 倍して加算する。そして、方向指数 2、4、6、8、10、12、14、および 16 に対応する各成分を削除して、16 値ベクトルから 8 値ベクトルへ圧縮する。残された 8 個の成分は、図4に示した 8 つの方向に対応している。

【0043】

次に、8 値ベクトルの成分のうち、互いに反対の向きに対応する 2 つの方向を同一視して、方向指数 1、3、5、および 7 に対応する各成分の値に、それぞれ方向指数 9、11、13、および 15 に対応する各成分の値を加算する。こうして、8 値ベクトルは 4 値ベクトルへ圧縮される。こうして、 $m \times n$ 個の 4 値ベク

トルからなる方向指数ヒストグラムベクトル場が得られる。

【0044】

ここで、横方向に並んだ n 個の 4 値ベクトルを方向指数ヒストグラム系列と呼ぶことにすると、 m 個の方向指数ヒストグラム系列が得られたことになる。特徴抽出部 11 は、これらの m 個の方向指数ヒストグラム系列を、画像の特徴量として出力する。

【0045】

例えば、図 8 のような空間情報圧縮に基づいて生成された方向指数ヒストグラムベクトル場の場合、図 9 に示すような 7 個の方向指数ヒストグラム系列が得られる。図 9 において、 a_{ij} ($i=1, 2, \dots, 7, j=1, 2, \dots, 13$) は 4 値ベクトルを表し、 v_i ($i=1, 2, \dots, 7$) は方向指数ヒストグラム系列を表す。

【0046】

次に、特徴生成部 13 による単語辞書の生成処理について説明する。今、単語リスト 14 として I 個の単語リスト $\delta_1, \delta_2, \dots, \delta_I$ が用意されており、 i 番目の単語リスト δ_i には、各単語の ID と、その単語を構成する各文字の ID のみが記されているものとする。ただし、各文字の ID は、個別文字辞書 15 にも登録されており、単語の特徴量を生成する際に参照されるものとする。

【0047】

特徴照合部 12 により処理対象の単語リスト δ_i が指定されると、特徴生成部 13 は、その中に含まれる各単語について、記されている構成文字の ID に基づいて個別文字辞書 15 を参照し、単語の特徴量を生成する。

【0048】

ここで、単語 w を構成する文字を c_1, c_2, \dots, c_K とし、 i 番目の文字 c_i の特徴量を Λ_i とする。各文字の特徴量は、あらかじめ上述した特徴量抽出と同様の処理により生成され、ID とともに個別文字辞書 15 に格納されている。このとき、単語 w の特徴量 Λ_w を、 $\Lambda_w = \sum \Lambda_i$ により定義する。ここで、 $\sum \Lambda_i$ は K 個の特徴量 $\Lambda_1, \Lambda_2, \dots, \Lambda_K$ の総和を表し、2 つの特徴量 Λ_1, Λ_2 の和 $\Lambda_1 + \Lambda_2$ は、次のような合成演算により定義される。

【0049】

今、 $\Lambda 1$ 、 $\Lambda 2$ を、それぞれ m 個の方向指数ヒストグラム系列 $\lambda 1 i$ 、 $\lambda 2 i$ ($i=1, 2, \dots, m$)を用いて、 $\Lambda 1 = (\lambda 1 1, \lambda 1 2, \dots, \lambda 1 m)$ 、 $\Lambda 2 = (\lambda 2 1, \lambda 2 2, \dots, \lambda 2 m)$ のように書くことにする。このとき、 $\Lambda 1 + \Lambda 2$ は、 m 個の方向指数ヒストグラム系列 $\lambda 1 i \lambda 2 i$ を用いて、 $(\lambda 1 1 \lambda 2 1, \lambda 1 2 \lambda 2 2, \dots, \lambda 1 m \lambda 2 m)$ と表される。

【0050】

ここで、 $\lambda 1 i \lambda 2 i$ は、方向指数ヒストグラム系列 $\lambda 1 i$ の後に方向指数ヒストグラム系列 $\lambda 2 i$ をそのまま付加して生成される新たな方向指数ヒストグラム系列を表す。 $\lambda 1 i$ 、 $\lambda 2 i$ がそれぞれ n 個の4値ベクトルからなる場合、 $\lambda 1 i \lambda 2 i$ は $2n$ 個の4値ベクトルからなる。

【0051】

例えば、 $m=7$ 、 $n=13$ とすると、特徴量 $\Lambda 1$ 、 $\Lambda 2$ 、および $\Lambda 1 + \Lambda 2$ は、図10のようになる。図10において、 $\Lambda 1$ の方向指数ヒストグラム系列 $\lambda 1 i$ ($i=1, 2, \dots, 7$)は、13個の4値ベクトル $a i j$ ($j=1, 2, \dots, 13$)からなり、 $\Lambda 2$ の方向指数ヒストグラム系列 $\lambda 2 i$ ($i=1, 2, \dots, 7$)は、13個の4値ベクトル $b i j$ ($j=1, 2, \dots, 13$)からなる。

【0052】

また、 $\Lambda 1 + \Lambda 2$ は $\Lambda 1$ と $\Lambda 2$ をそのまま横方向に並べて生成され、その方向指数ヒストグラム系列 $\lambda 1 i \lambda 2 i$ ($i=1, 2, \dots, 7$)は、26個の4値ベクトル $c i j$ ($j=1, 2, \dots, 26$)からなる。 $c i 1 \sim c i 13$ は $a i 1 \sim a i 13$ に一致し、 $c i 14 \sim c i 26$ は $b i 1 \sim b i 13$ に一致する。言い換えれば、 $j=1, 2, \dots, 13$ については $c i j = a i j$ であり、 $j=14, 15, \dots, 26$ については $c i j = b i (j-13)$ である。

【0053】

次に、図11から図13までを参照しながら、特徴照合部12の処理について説明する。図11は、特徴照合部12および特徴生成部13の処理のフローチャートである。ここでは、照合に用いられる単語リスト14には、0番目からS-

1 番目までの S 個の単語が含まれているものとする。

【0054】

まず、特徴照合部 12 は、制御変数 i を初期値 0 に設定し（ステップ S11）、単語リスト 14 に含まれる単語の総数 S と i を比較する（ステップ S12）。そして、 i が S より小さければ、特徴生成部 13 に i 番目の単語の特徴量を生成するよう依頼する。

【0055】

この依頼を受けて、特徴生成部 13 は、単語リスト 14 にアクセスし（ステップ S13）、上述したような処理を行って、個別文字辞書の特徴量から i 番目の単語の特徴量を生成する（ステップ S14）。そして、生成された単語の特徴量を特徴照合部 12 に出力する。

【0056】

次に、特徴照合部 12 は、特徴抽出部 11 から入力された画像の特徴量と特徴照合部 12 から入力された i 番目の単語の特徴量をメモリ上で照合し、2 つの特徴量の距離（類似度）を計算する（ステップ S15）。

【0057】

次に、 i 番目の単語の特徴量が保持されているメモリ領域を解放し（ステップ S16）、 i を 1 だけインクリメントして（ステップ S17）、ステップ S12 以降の処理を繰り返す。ステップ S16 でメモリ領域がフリーになっているため、 $i+1$ 番目の単語の特徴量を同じ領域に書き込むことができ、使用するメモリ量を節約できる。そして、ステップ S12 において i が S に達すると、処理を終了する。

【0058】

ステップ S15 において、特徴照合部 12 は、横方向の文字のずれや変形を吸収するために、次のような距離計算を行う。まず、入力画像の特徴量を $N = (v_1, v_2, \dots, v_m)$ とし、比較対象となる単語の特徴量を $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$ とする。ただし、 v_i, λ_i ($i=1, 2, \dots, m$) は、図 9 に示したような方向指数ヒストグラム系列である。このとき、2 つの特徴量 N と Λ の距離 $D(N, \Lambda)$ は、

$$D(N, \Lambda) = \sum D(v_i, \lambda_i) \quad (1)$$

と表される。ここで、 $\sum D(v_i, \lambda_i)$ は、2つの方向指数ヒストグラム系列 v_i と λ_i の距離 $D(v_i, \lambda_i)$ の i に関する総和である。

【0059】

上述したように、方向指数ヒストグラム系列は4値ベクトルの系列であるから、それをベクトルの成分毎に分解して考えると、方向指数ヒストグラム系列は4つの数値系列とみなすことができる。そこで、方向指数ヒストグラム系列 v_i の j 番目の数値系列を $v_i(j)$ ($j=1, 2, 3, 4$) と書くことにすると、

$$\begin{aligned} v_i &= (v_i(1), v_i(2), v_i(3), v_i(4)), \\ \lambda_i &= (\lambda_i(1), \lambda_i(2), \lambda_i(3), \lambda_i(4)) \end{aligned} \quad (2)$$

と表される。このとき、 $D(v_i, \lambda_i)$ は、

$$D(v_i, \lambda_i) = \sum D(v_i(j), \lambda_i(j)) \quad (3)$$

と表される。ここで、 $\sum D(v_i(j), \lambda_i(j))$ は、2つの数値系列 $v_i(j)$ と $\lambda_i(j)$ の距離 $D(v_i(j), \lambda_i(j))$ の j に関する総和である。そして、距離 $D(v_i(j), \lambda_i(j))$ は、動的計画法 (Dynamic Programming, DP) により計算することができる。

【0060】

DPマッチングは、音声等の時系列データのマッチング方法としてよく知られている。この方法では、2つのデータの集合を照合する際に、データの局所的な特徴に注目して、全体的な対応付けの良否を表す評価関数が定義される。ここでは、この評価関数の値から2つのデータの距離が計算される。

【0061】

図12は、 n 個の数値からなる数値系列 $\{x_1, x_2, \dots, x_n\}$ と、 p 個の数値からなる数値系列 $\{y_1, y_2, \dots, y_p\}$ のDPマッチングの方法を示している。

【0062】

ここでは、 xy 座標平面において、 x 軸上に数値系列 $\{x_1, x_2, \dots, x_n\}$ を並べ、 y 軸上に数値系列 $\{y_1, y_2, \dots, y_p\}$ を並べて、2つの数値系列の対応関係を平面上の複数の点で表す。そして、点 (x_1, y_1) を

始点として、計算領域A内であらかじめ決められた漸化式に従って評価関数 $g(x_i, y_j)$ を順に計算しながら、2つの数値系列の点同士を対応付けていく。そして、最終的に得られた $g(x_n, y_p)$ から2つの数値系列の距離を求める。

【0063】

図13は、DPマッチングの過程において、既に求められた $g(x_{i-1}, y_j)$ 、 $g(x_{i-1}, y_{j-1})$ 、および $g(x_{i-1}, y_{j-1})$ から $g(x_i, y_j)$ を求める計算を示している。ここでは、例えば、次のような漸化式が用いられる。

$$\begin{aligned} g(x_i, y_j) \\ = \min \{ & g(x_{i-1}, y_j) + d(x_{i-1}, y_j), \\ & g(x_{i-1}, y_{j-1}) + 2 * d(x_{i-1}, y_{j-1}), \\ & g(x_i, y_{j-1}) + d(x_i, y_{j-1}) \} \end{aligned} \quad (4)$$

ここで、 $g(x_i, y_j)$ は、部分数値系列 $\{x_1, x_2, \dots, x_i\}$ を部分数値系列 $\{y_1, y_2, \dots, y_j\}$ に対応付けた時点における評価関数の値を表す。また、 $d(x_i, y_j)$ は、数値 x_i と数値 y_j を対応付けたときの距離を表し、例えば、次式により与えられる。

$$d(x_i, y_j) = |x_i - y_j| \quad (5)$$

また、 $\min \{ \}$ は $\{ \}$ 内の3つの要素のうちの最小値を表す。こうして、(4)式により、 $g(x_i, y_j)$ を最小にするような部分数値系列 $\{x_1, x_2, \dots, x_i\}$ と $\{y_1, y_2, \dots, y_j\}$ の対応付けが採用され、 $g(x_i, y_j)$ が記憶される。

【0064】

このような計算を繰り返すことで、最終的に数値系列 $\{x_1, x_2, \dots, x_n\}$ と $\{y_1, y_2, \dots, y_p\}$ が対応付けられ、 $g(x_n, y_p)$ が求められる。そして、 $g(x_n, y_p) / (n+p)$ を2つの数値系列の間の距離とする。距離が小さいほど、2つの数値系列は類似しており、距離が大きいほど、2つの数値系列は相異している。

【0065】

こうして、DPマッチングにより距離 $D(v_i(j), \lambda_i(j))$ を計算すれば、(1)、(3)式から2つの特徴量の距離 $D(N, \Lambda)$ を計算することができる。

【0066】

DPマッチングにおいては、2つの数値の組み合わせの生成方法に柔軟性があり、2つの数値系列が非線形に対応付けられる。この柔軟性を用いて画像の横方向の特徴のずれをある程度吸収することができる。このように、本実施形態では、従来のぼかし処理の効果を特徴量の距離計算におけるDPマッチングで代替することで、文字の連結方向のぼかし処理を施さない特徴量の使用を可能にしている。なお、特徴量の距離計算としては、DPマッチングのほかにも、特徴のずれを吸収できるような任意の非線形マッチング方法を用いることができる。

【0067】

次に、入力画像の具体例を用いて処理のフローを説明する。今、図14のような文字列画像のうち、“川崎市”の部分の認識処理が終了し、続いて“中原”の部分が入力された場合を考える。このとき、“中原”の入力画像から特徴量が抽出され、図15のようなフローに従って、特徴照合処理が行われる。

【0068】

単語リスト14には、“川崎”、“幸”、“中原”、“高津”、“宮前”、“多摩”、および“麻生”が単語として登録されており、個別文字辞書15には、すべての文字の特徴量が登録されているものとする。

【0069】

まず、単語リスト14に従って、個別文字辞書15から“川”と“崎”の特徴量が取り出され、“川崎”の特徴量が合成される(ステップS21)。次に、合成された“川崎”の特徴量と入力画像の特徴量が照合され、2つの特徴量の距離が記憶される(ステップS22)。そして、“川崎”の特徴量のメモリ領域が解放される(ステップS23)。

【0070】

次に、“幸”は1文字単語であるので、個別文字辞書15から“幸”の特徴量が取り出され、解放されたメモリ領域に格納される(ステップS24)。そして

、“幸”の特徴量と入力画像の特徴量が照合されて、2つの特徴量の距離が記憶され（ステップS25）、“幸”の特徴量のメモリ領域が解放される（ステップS26）。

【0071】

次に、個別文字辞書15から“中”と“原”の特徴量を取り出され、解放されたメモリ領域で“中原”の特徴量が合成される（ステップS27）。そして、合成された“中原”の特徴量と入力画像の特徴量が照合されて、2つの特徴量の距離が記憶され（ステップS28）、“中原”の特徴量のメモリ領域が解放される（ステップS29）。

【0072】

このような照合処理が単語リスト14のすべての単語について行われ、照合処理が終了すると、それらの単語が距離の小さいものから順に認識結果として出力される。

【0073】

図14、15では、漢字からなる単語の処理が記載されているが、平仮名、片仮名、英数字、記号等を含む単語についても同様の処理が行われる。また、日本語のみならず、中国語、韓国語、英語、ドイツ語、フランス語等の任意の言語における単語認識にも同様の処理が適用される。

【0074】

さらに、本発明は、単語認識という枠組みを越えて、1つ以上の個別パターンからなるパターン列の認識処理にも適用できる。この場合、単語リストの代わりに認識結果の候補となるパターン列を登録したリストを用意し、個別文字辞書の代わりに個別パターンの特徴量を登録した辞書を用意しておく。そして、画像の認識処理の過程で、個別パターンの辞書から動的にパターン列の特徴量を生成し、パターン列を一括認識する。

【0075】

ところで、図2の単語認識装置は、図16に示すような情報処理装置（コンピュータ）を用いて構成することができる。図16の情報処理装置は、CPU（中央処理装置）21、メモリ22、入力装置23、出力装置24、外部記憶装置2

5、媒体駆動装置 26、ネットワーク接続装置 27、および光電変換装置 28 を備え、それらはバス 29 により互いに接続されている。

【0076】

メモリ 22 は、例えば、ROM (read only memory)、RAM (random access memory) 等を含み、処理に用いられるプログラムとデータを格納する。CPU 21 は、メモリ 22 を利用してプログラムを実行することにより、必要な処理を行う。

【0077】

図 2 の特徴抽出部 11、特徴照合部 12、および特徴生成部 13 は、メモリ 22 の特定のプログラムコードセグメントに格納されたソフトウェアコンポーネントに対応する。また、単語リスト 14 および個別文字辞書 15 は、データとしてメモリ 22 の特定の領域に格納される。

【0078】

入力装置 23 は、例えば、キーボード、ポインティングデバイス、タッチパネル等であり、ユーザからの指示や情報の入力に用いられる。出力装置 24 は、例えば、ディスプレイ、プリンタ、スピーカ等であり、ユーザへの問い合わせや情報の出力に用いられる。

【0079】

外部記憶装置 25 は、例えば、磁気ディスク装置、光ディスク装置、光磁気ディスク (magneto-optical disk) 装置等であり、取得された情報を格納する。この外部記憶装置 25 に、上述のプログラムとデータを保存しておき、必要に応じて、それらをメモリ 22 にロードして使用することもできる。

【0080】

媒体駆動装置 26 は、可搬記録媒体 30 を駆動し、その記録内容にアクセスする。可搬記録媒体 30 としては、メモリカード、フロッピーディスク、CD-ROM (compact disk read only memory)、光ディスク、光磁気ディスク等、任意のコンピュータ読み取り可能な記録媒体が用いられる。この可搬記録媒体 30 に上述のプログラムとデータを格納しておき、必要に応じて、それらをメモリ 22 にロードして使用することもできる。

【0081】

ネットワーク接続装置 27 は、LAN (local area network) 等の任意のネットワーク (回線) を介して外部の装置と通信し、通信に伴うデータ変換を行う。また、必要に応じて、上述のプログラムとデータを外部の装置から受け取り、それらをメモリ 22 にロードして使用することもできる。

【0082】

光電変換装置 28 は、例えば、イメージスキャナ等であり、認識対象の画像をデジタルデータに変換して入力する。入力された画像データはメモリ 22 に読み込まれ、そのデータから特徴量が抽出される。

【0083】

図 17 は、図 16 の情報処理装置にプログラムとデータを供給することのできるコンピュータ読み取り可能な記録媒体を示している。可搬記録媒体 30 や外部のデータベース 31 に保存されたプログラムとデータは、メモリ 22 にロードされる。そして、CPU 21 は、そのデータを用いてそのプログラムを実行し、必要な処理を行う。

【0084】

【発明の効果】

本発明によれば、個別文字辞書から動的に単語辞書を作成することにより、単語の範囲を制限することなく、単語を一括して認識することができる。したがって、任意の用途における単語認識が可能になる。

【0085】

また、本発明では、特徴量の距離計算に DP マッチングのような非線形なマッチング方法を採用し、文字の連結方向におけるメッシュの区切り方を可変にしているので、単語認識においても一定の認識精度が保たれる。

【図面の簡単な説明】

【図 1】

本発明の単語認識装置の原理図である。

【図 2】

単語認識装置の構成図である。

【図 3】

特徴抽出部の処理のフローチャートである。

【図 4】

輪郭点の位置と方向指数の関係を示す図である。

【図 5】

16 方向の方向指数を示す図である。

【図 6】

16 方向の方向指数の決定方法を示す図である。

【図 7】

一次元ガウス分布型フィルタを示す図である。

【図 8】

基本メッシュ分割とマスクの中心位置を示す図である。

【図 9】

方向指数ヒストグラム系列を示す図である。

【図 10】

特徴量の合成を示す図である。

【図 11】

特徴照合部および特徴生成部の処理のフローチャートである。

【図 12】

DP マッチングを示す図である。

【図 13】

DP マッチングの過程を示す図である。

【図 14】

画像の例を示す図である。

【図 15】

特徴照合処理の例を示す図である。

【図 16】

情報処理装置の構成図である。

【図 17】

記録媒体を示す図である。

【符号の説明】

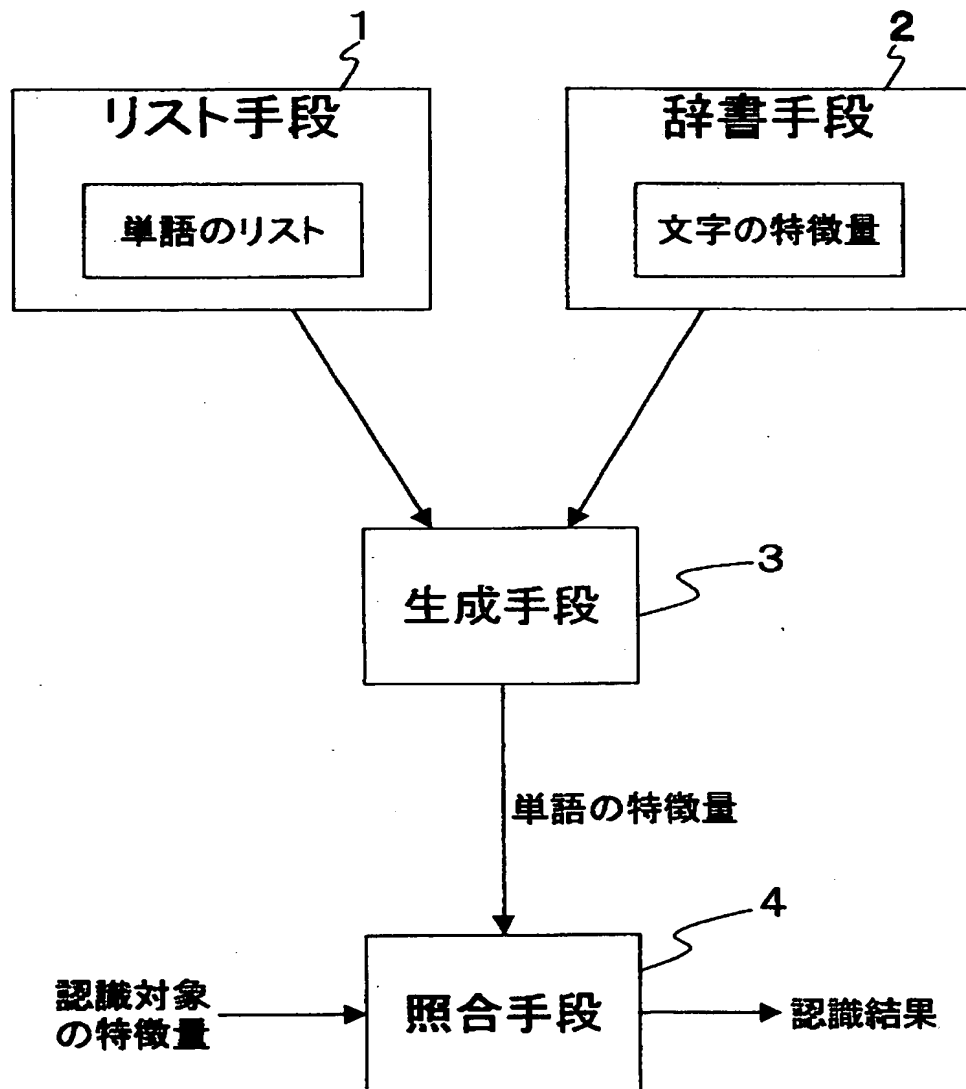
- 1 リスト手段
- 2 辞書手段
- 3 生成手段
- 4 照合手段
- 11 特徴抽出部
- 12 特徴照合部
- 13 特徴生成部
- 14 単語リスト
- 15 個別文字辞書
- 21 CPU
- 23 入力装置
- 24 出力装置
- 25 外部記憶装置
- 26 媒体駆動装置
- 27 ネットワーク接続装置
- 28 光電変換装置
- 29 バス
- 30 可搬記録媒体
- 31 データベース

特平 10-308943

【書類名】 図面

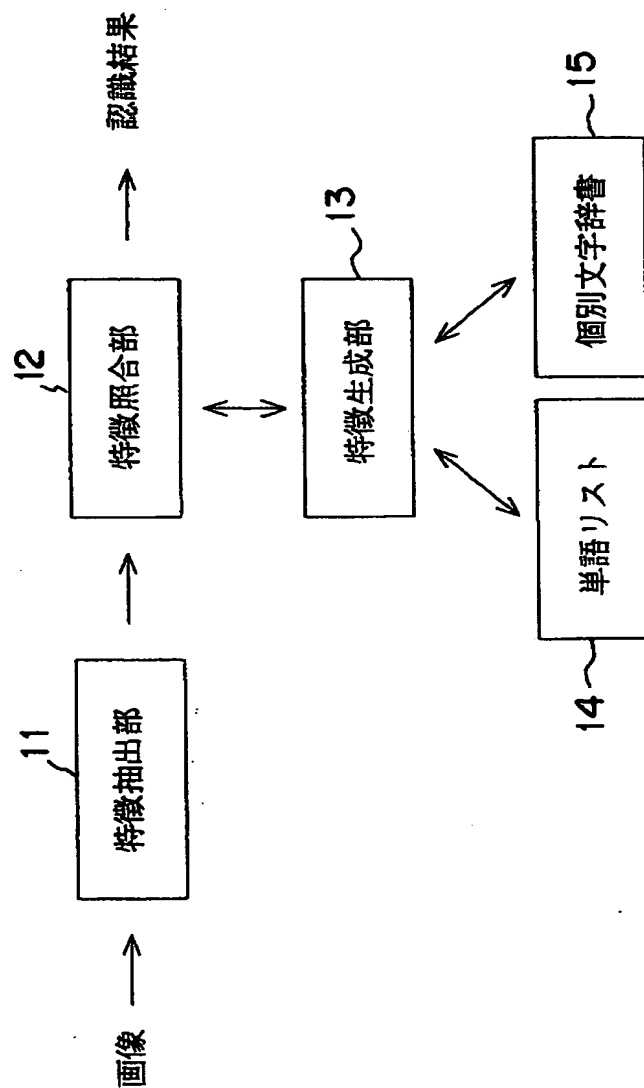
【図 1】

本発明の単語認識装置の原理図



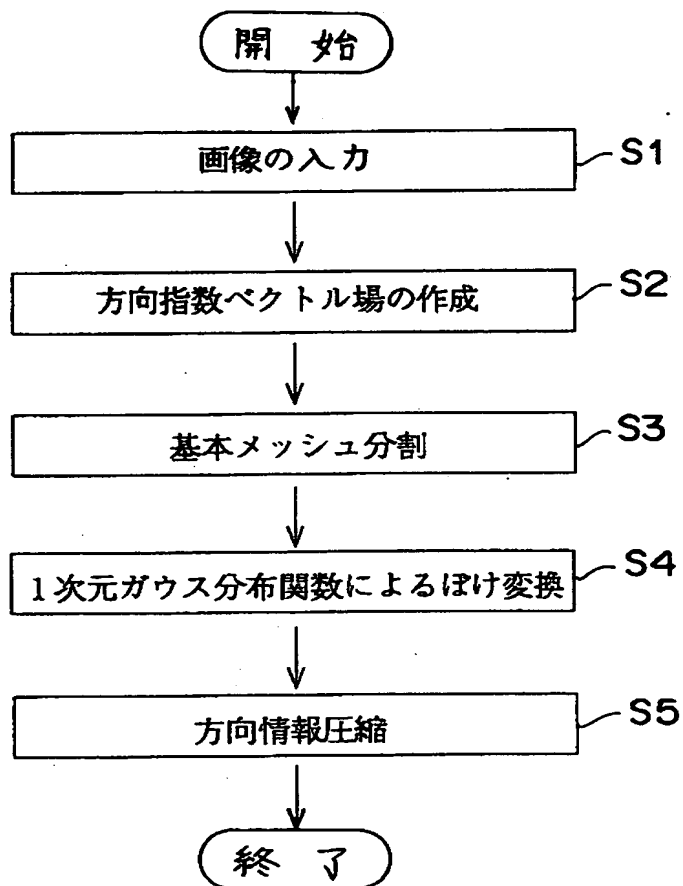
【図 2】

単語認識装置の構成図



【図 3】

特徴抽出部の処理のフローチャート



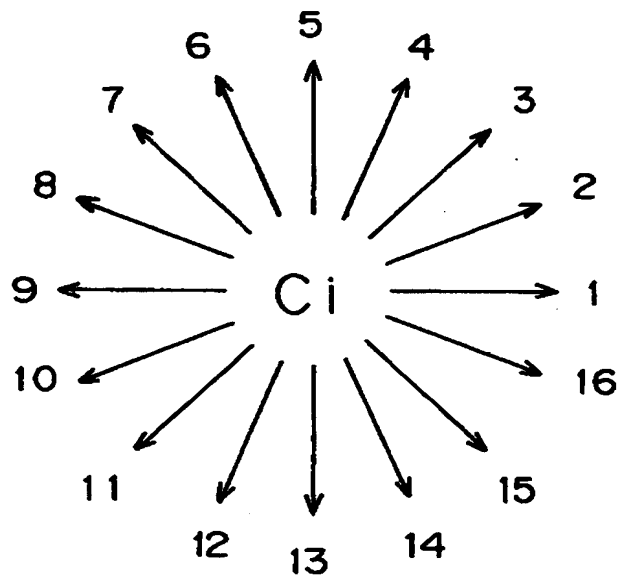
【図4】

輪郭点の位置と方向指数の
関係を示す図

7	5	3
9	Ci	1
11	13	15

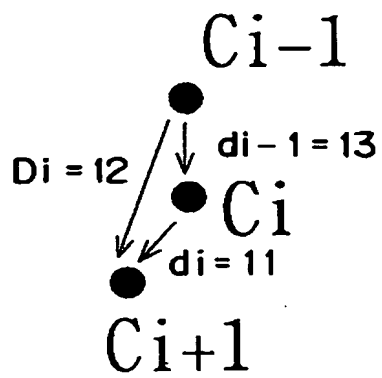
【図5】

16方向の方向指数を示す図



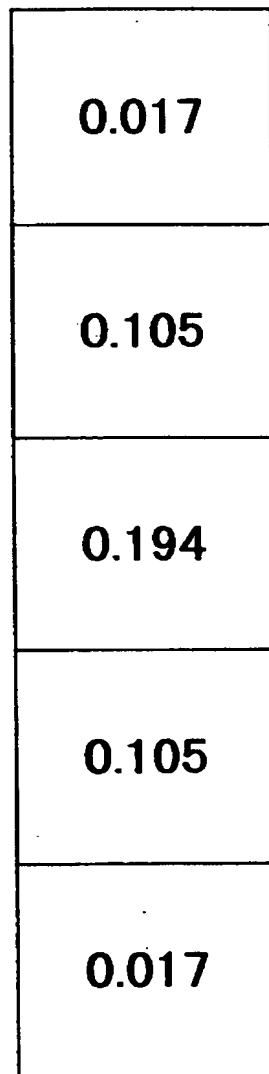
【図 6】

16 方向の方向指数の決定方法
を示す図



【図 7】

一次元ガウス分布型フィルタを示す図



【図8】

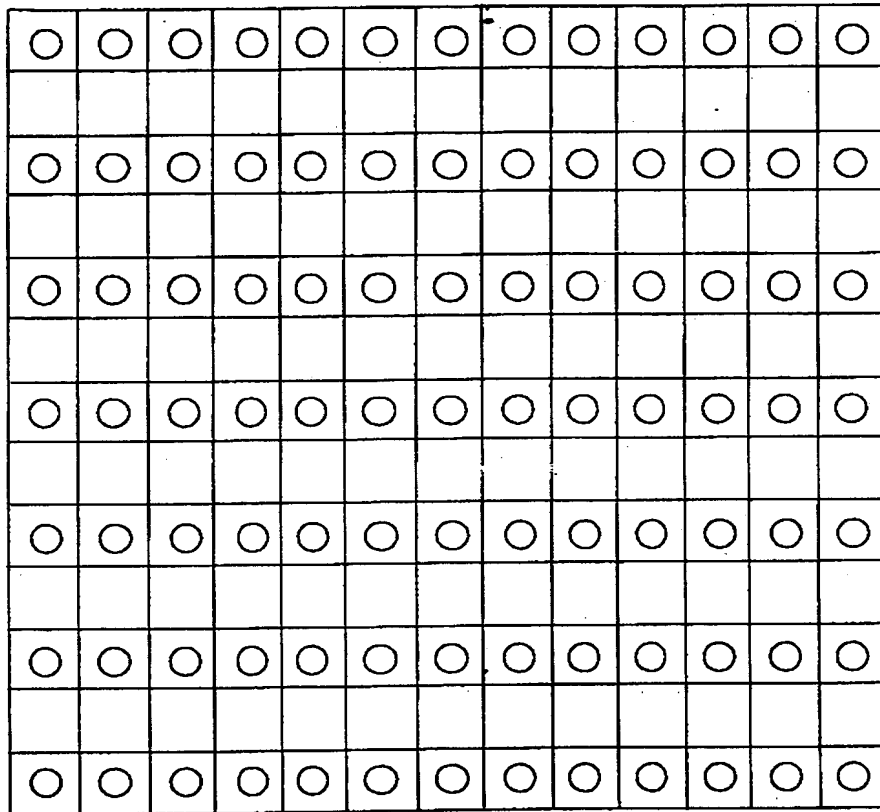
基本メッシュ分割と

マスクの中心位置を示す図

$n = 13$

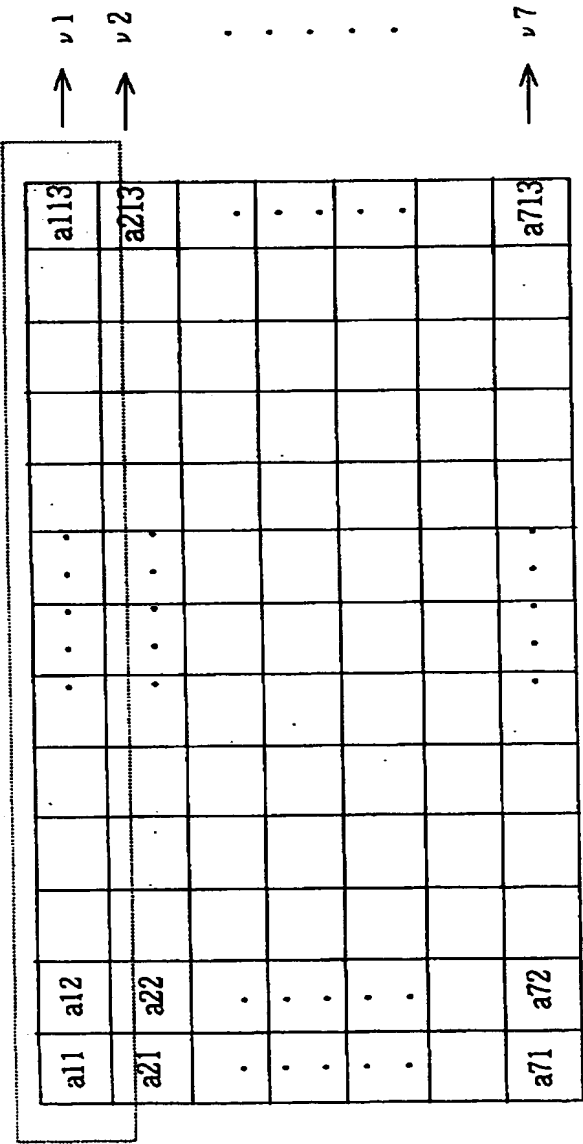
$M = 13$

$m = 7$



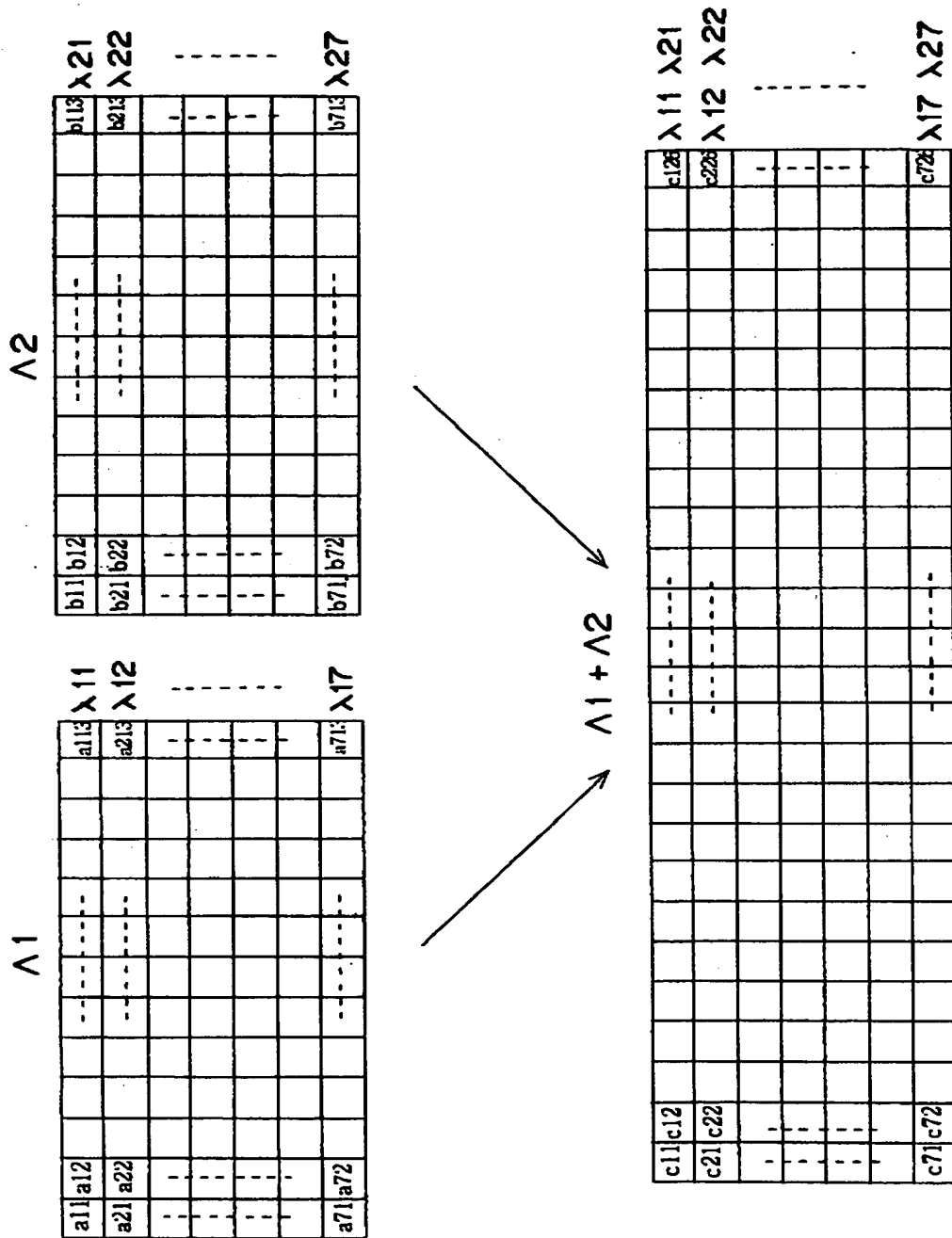
【図9】

方向指数ヒストグラム系列を示す図



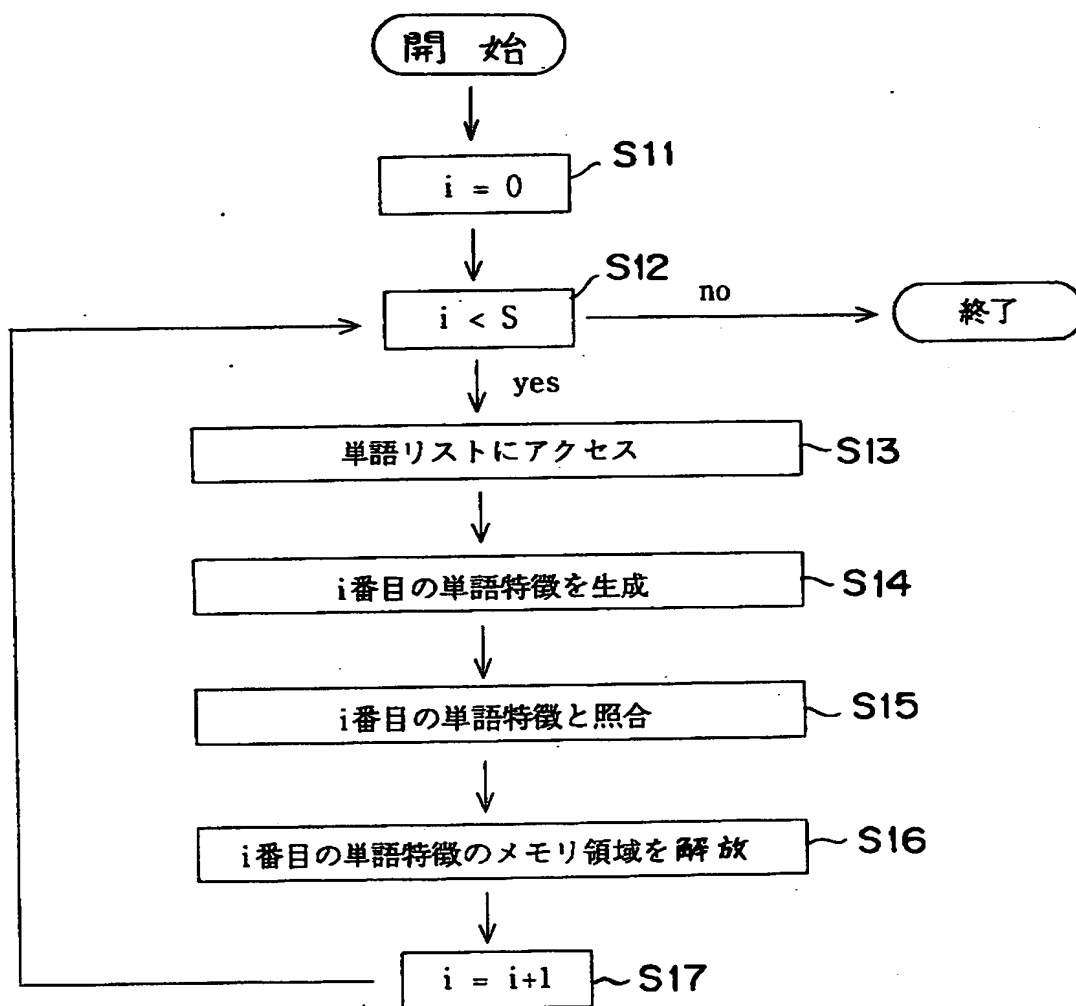
【図10】

特徴量の合成を示す図



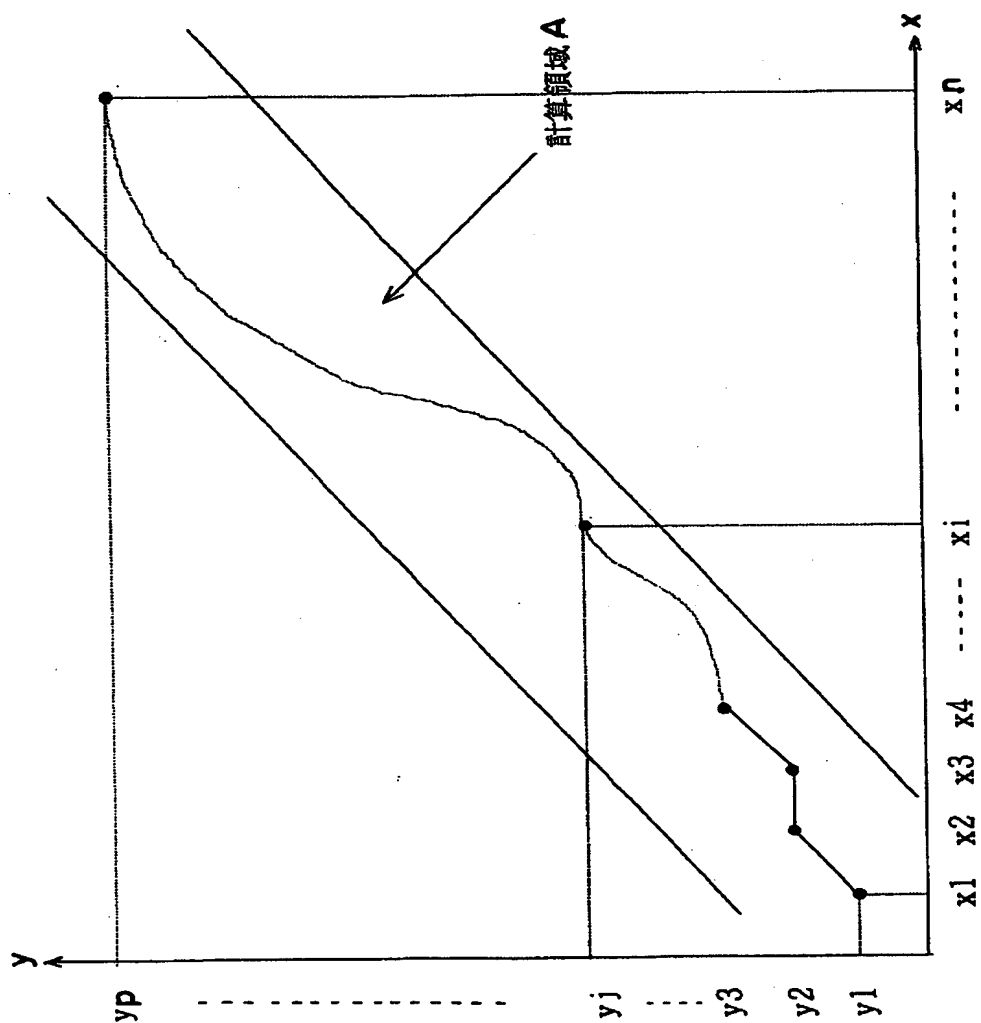
【図 11】

特徴照合部 および特徴生成部の
処理のフローチャート



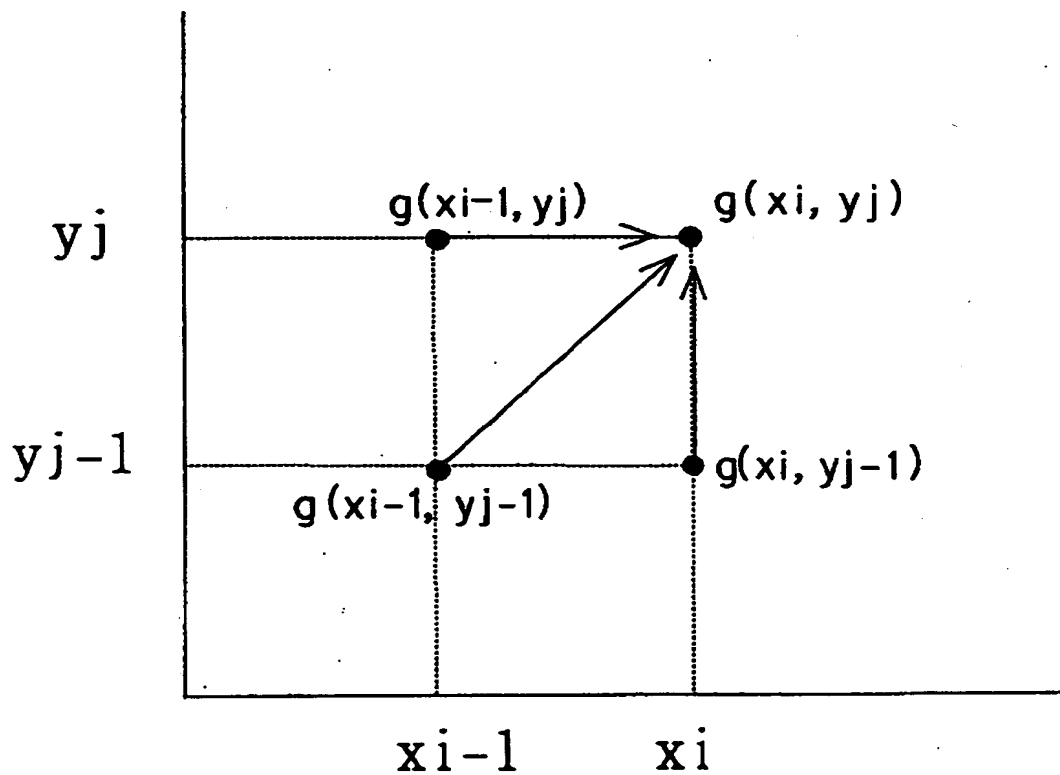
【図 12】

DP マッチングを示す図



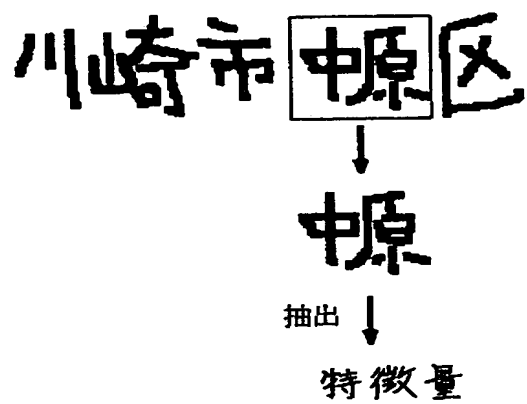
【図 13】

DP マッチングの過程を示す図



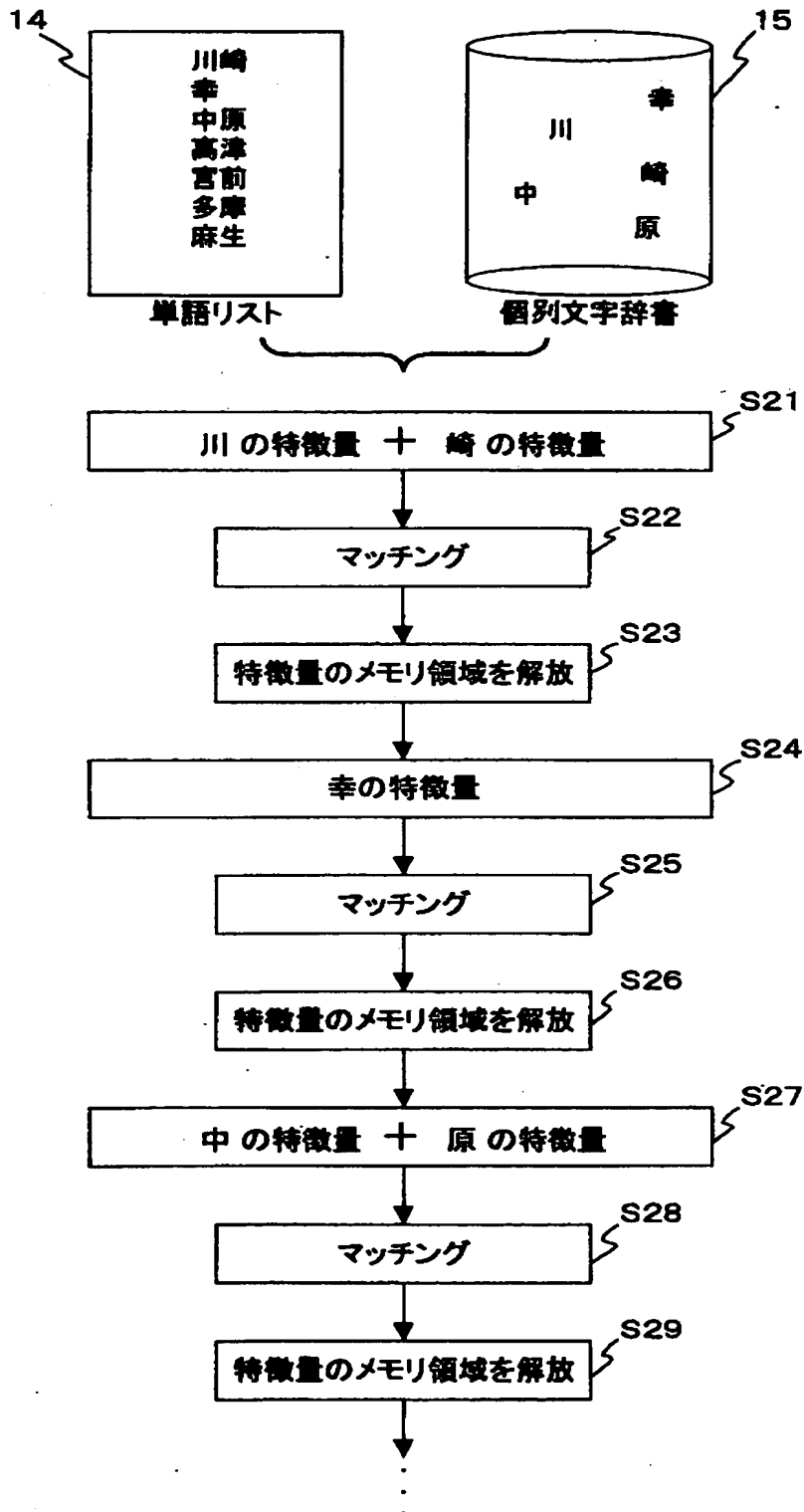
【図14】

画像の例を示す図



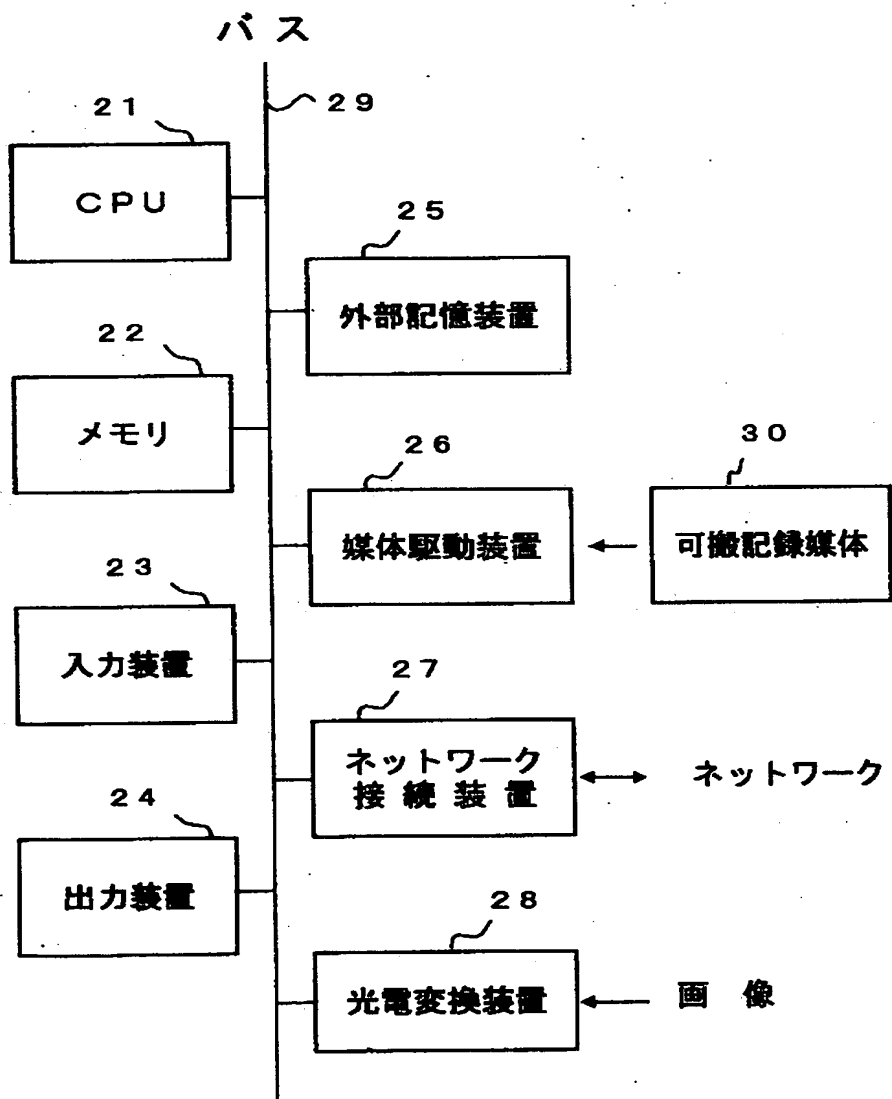
【図15】

特徴照合処理の例を示す図



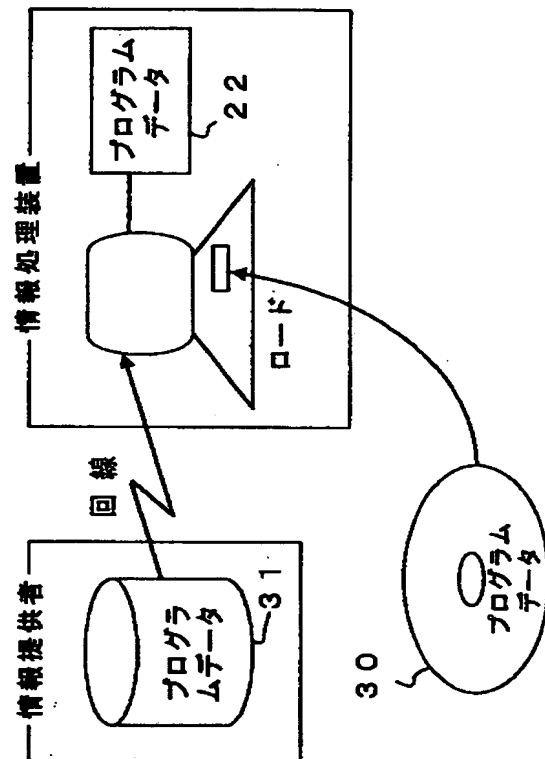
【図 16】

情報処理装置の構成図



【図 17】

記録媒体を示す図



【書類名】 要約書

【要約】

【課題】 認識可能な単語の範囲をできるだけ制限することなく、単語を一括して認識することが課題である。

【解決手段】 特徴抽出部 11 は、与えられた画像から特徴量を抽出し、特徴生成部 13 は、個別文字辞書 15 に登録された文字の特徴量を用いて、単語リスト 14 に登録された認識候補の単語の特徴量を動的に合成する。特徴照合部 12 は、合成された単語の特徴量を画像から抽出された特徴量と照合し、2 つの特徴量の類似度を計算して、認識結果を出力する。

【選択図】 図 2

【書類名】 職権訂正データ
【訂正書類】 特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 000005223

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号

【氏名又は名称】 富士通株式会社

【代理人】 申請人

【識別番号】 100074099

【住所又は居所】 東京都千代田区二番町8番地20 二番町ビル3F
大菅内外国特許事務所

【氏名又は名称】 大菅 義之

【選任した代理人】

【識別番号】 100067987

【住所又は居所】 神奈川県横浜市港北区太尾町1418-305 (大倉山二番館) 久木元特許事務所

【氏名又は名称】 久木元 彰

出 願 人 履 歴 情 報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号
氏 名 富士通株式会社